



Wang, D., Canagarajah, CN., Redmill, DW., & Bull, DR. (2004). Multiple description video coding based on zero padding. In *2004 IEEE International Symposium on Circuits and Systems (ISCAS '04) Vancouver, BC, Canada* (Vol. 2, pp. 205 - 208). Institute of Electrical and Electronics Engineers (IEEE).  
<https://doi.org/10.1109/ISCAS.2004.1329244>

Peer reviewed version

Link to published version (if available):  
[10.1109/ISCAS.2004.1329244](https://doi.org/10.1109/ISCAS.2004.1329244)

[Link to publication record in Explore Bristol Research](#)  
PDF-document

## University of Bristol - Explore Bristol Research

### General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:  
<http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>

# MULTIPLE DESCRIPTION VIDEO CODING BASED ON ZERO PADDING

*D. Wang<sup>†</sup>, N. Canagarajah, D. Redmill and D. Bull*

Centre for Communications Research,  
University of Bristol, Bristol, BS8 1UB, UK  
Tel: +44 (0)117 9545202, Fax: +44 (0)117 9545206  
<sup>†</sup>Email: Dong.Wang@bristol.ac.uk

## ABSTRACT

This paper proposes a simple multiple description video coding approach based on zero padding theory. It's completely based on pre- and post-processing, which requires no modifications to the source codec. Redundancy is added by padding zeros in DCT domain, which results in interpolation of original frame and increases correlations between pixels. Methods based on 1D and 2D DCT are presented. We also investigate two sub-sampling method, which are interleaved and quincunx, to generate multiple descriptions. Results are presented for two zero padding approaches using H.264, which shows that the 1D approach performs much better than 2D padding techniques, at a much lower computational complexity. For 1D zero padding, results show that interleaved sub-sampling is better than quincunx.

## 1. INTRODUCTION

Video transmission over lossy network is always a big challenge. To video, due to predictive coding, any bit loss may cause great quality degradation. Multiple description coding is one approach to address this problem. It generates several sub bit streams called descriptions from source video instead of one. Each description can reconstruct video of acceptable quality and all the descriptions together can reconstruct high quality video.

Unlike layered video coding techniques, each description generated by MDC can independently be decoded and reconstructed to acceptable quality. This can give a graceful degradation of received video with loss, while it also avoid catastrophic failure of layered coding due to loss of base layer.

There are basically two kinds of decoders in MDC system. One is called central decoder which is used when all the descriptions are received. The other is side decoder which just uses one or a subset of descriptions to reconstruct video of acceptable quality.

It's clear that more correlations in descriptions will result in higher quality of side decoded video, for fewer information is lost. But at the same time central decoder must perform with lower efficiency because more redundancy is introduced. Extensive research on MDC to increase the efficiency has been conducted.

MDC based on Scalar Quantization is developed in [1] to divide a signal by two coarser quantizers. Output of each

quantizer is the approximation of single description. Any description can use its coarse data to generate a basic video and both of them can be combined to reconstruct high quality video. It's applied on predictive video coding in [2]. Another approach on image coding is addressed in [3] and [4] using pairwise correlating transforms to transform a vector of DCT coefficients into another vector of correlated components, which introduces additional redundancy between components of vector. It's used in [5] for the problem of motion compensated video coding. In [6], video sequence is divided into two by means of odd and even frames and different concealment methods are used to estimate lost frames. Overlapping technique is used on motion vectors in [7] to achieve more accurate prediction of lost data.

Another simple way of generating MDC is that through pre- and post-processing, as in [8] and [9]. Here we name it as Two Dimensional Multiple Description Coding (2D-MDC). The technique is based on the theory of zero padding. Padding zeros in time domain can result in interpolation in frequency domain. On the other hand, padding zeros in frequency domain results in interpolation in time domain. The source video frames are transformed using two dimensional DCT (2-D DCT). In DCT domain, certain number of zeros is padded in both horizontal and vertical directions. Coefficients of new size are inverse transformed using 2D IDCT and results in an over-sampled frame of bigger size. Pixels are more correlated now and then video source is divided into two descriptions by sub-sampling. The two descriptions are independently coded at the encoder. If one description is lost, it can be estimated by the other one. This proved to be much better than directly sub-sampling the original frame although it increases the image size to be coded. It performs very well especially at low bit rate, and it's very simple to implement, since no modification is required on the source coder. But there exist some problems of this proposed method. Firstly, it adds lots of zeros in both directions which results in much bigger size and lower central efficiency. Secondly, it just uses nearest pixel in the same column to conceal the lost sample in the other description. This doesn't utilize the added correlations sufficiently, for it only estimate sample by one dimension, while additional correlations also exist in horizontal direction. Our experiments show that estimation using adjacent four pixels gives improvement on this way. Moreover, two dimensional transform brings relatively high computational complexity.

We propose a new method based on zero padding by 1-D DCT (here named 1D-MDC). This technique was considered for

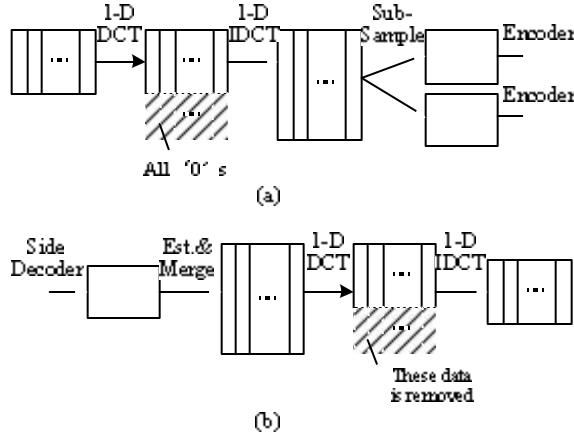


image coding in [12] and provided good result. In this paper we develop a video coding scheme based on this 1-D DCT approach. The proposed scheme has less computational complexity and performs better than the 2D-MDC. Performance using two sub-sampling methods based on latest video coding standard H.264 is evaluated to provide a suitable comparison.

The rest of this paper is organized as follows. In Section 2 1D-MDC with different sub-sampling and estimation methods are described. Section 3 gives the results and analysis of experiments and compares our method with 2D-MDC. Conclusions are presented in section 4.

## 2. DESCRIPTION OF THE PROPOSED 1D-MDC

The basic structure of zero padding method is shown in Fig 1 (a). Firstly, each frame is transformed using 1-D DCT on each column, then padded with zeros vertically. After 1-D IDCT on new-sized column it's sub-sampled into two sub frames and independently coded. As mentioned in section 1, after zero padding pixels are more correlated and can lead to higher side decoding quality when one description is lost.

In Central decoder the two reconstructed sub frames are merged together. Then DCT, removing padded zeros and IDCT is performed to get the frame needed.

When only one description is received, as in Fig. 1 (b), the other description is estimated using correctly received description and then merged together, like in central decoder, to get a concealed frame.

There are two contradictory factors in zero padding technique for central quality. The first is that picture size is enlarged compared to no zero padded, and it results in more macroblocks to be encoded at the encoder. The second is that with padding more zeros, correlations between pixels are increased and pictures are smoother. There will be less high frequency in picture and it results in fewer bits used for one macroblock. Results show that encoding additional macroblocks is less important to total bitrate than higher correlations, and the second factor is dominant for encoding. We will see from the next section that adding certain more zeros makes the result better.

It's clear that number of padded zeros also affect side decoding quality. When more zeros are padded, correlation between two descriptions will be higher producing better estimation. If central quality doesn't drop much, adding more zeros means better performance. But with central quality decreasing, a tradeoff should be considered between central and side quality.

It's worth noting that the sub-sampling and estimation method play an important role in this approach. In this work, we consider various sub sampling schemes, including Quincunx for in [9] this sub sampling method is used. Pixels in (even row, even column) and (odd row, odd column) belongs to one description and remaining belongs to the other. But the problem in [9] is that estimation just uses nearest pixels in the same column. Added correlation is not utilized sufficiently, for it only uses vertical correlation while it adds zeros in both directions. The improvement which uses mean value of adjacent four pixels will be shown for comparison in the next section.

Interleaved sub-sampling method, which only utilizes vertical correlation, is suitable in 1D-MDC, for the added redundancy only exists in the vertical direction. It's simplest to implement. Quincunx sub-sampling method is also suitable to be used in 1D-MDC, for there is original redundancy existing in picture. Two estimation methods may be used based on quincunx sub sampling. One is that just uses adjacent pixels in same column to estimate lost pixel, for more redundancy is in vertical direction. The other is using mean value of adjacent four pixels, but required to be modified to weighted mean value. Lost pixels  $x$  is estimated by adjacent four pixels,  $a$  and  $c$  for same column,  $b$  and  $d$  for same row, using

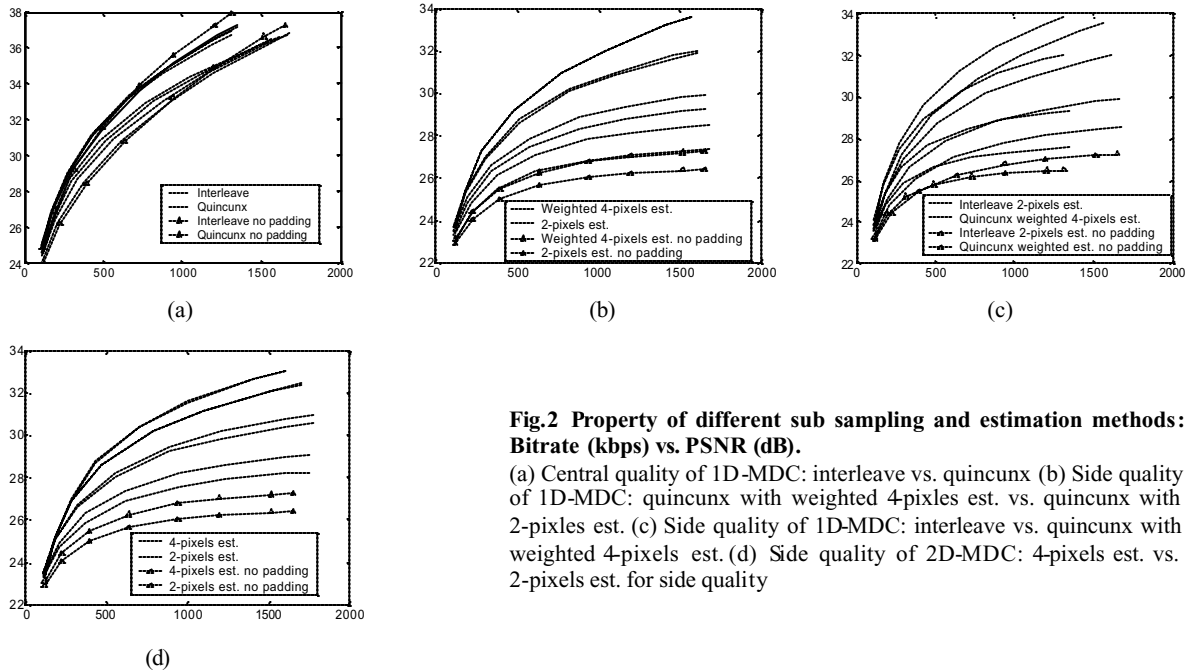
$$x = w \times \frac{(a+c)}{2} + (1-w) \times \frac{(b+d)}{2} \quad (1)$$

where  $w$  is a weighting parameter. The reason is that we add correlation in the vertical direction and so vertical correlation should be larger than horizontal. This method utilizes all four adjacent pixels for estimation and seems to be of the best performance in above approaches. But results show that, although it's better than two dimensional zero padding technique, it's not the best because of higher encoding bit rate.

Another thing needed to be considered is that, after IDCT at pre-processing stage, most of pixels must be non-integer, while the standard coder needs it to be integer. So the fraction part will be dropped and this leads to a small mismatch. With increasing zeros, this mismatch will also be bigger. Fortunately, in all video coder, quantization is performed and this mitigates the effect of this mismatch with certain number of zeros padded.

## 3. RESULTS AND ANALYSIS

We examine the performance of our proposed 1D-MDC compared to 2D-MDC. All the comparisons are made by assuming that one entire description is lost. For 1D-MDC, two sub sampling method are chosen. One is interleaved sub-sampling which just takes odd row as one description and even row as the other. For this sampling method, when doing estimation, we use the mean value of adjacent two pixels in the same column to conceal the lost pixel (here called Interleave 2-pixels est.). The other is quincunx sub-sample mentioned above. Two estimation methods can be used of which one is by 2-pixels est. (Quincunx 2-pixels est.) and the other is by adjacent 4 pixels (Quincunx weighted 4 pixels est.). For 2D-MDC, just



**Fig.2 Property of different sub sampling and estimation methods: Bitrate (kbps) vs. PSNR (dB).**

(a) Central quality of 1D-MDC: interleave vs. quincunx (b) Side quality of 1D-MDC: quincunx with weighted 4-pixels est. vs. quincunx with 2-pixels est. (c) Side quality of 1D-MDC: interleave vs. quincunx with weighted 4-pixels est. (d) Side quality of 2D-MDC: 4-pixels est. vs. 2-pixels est. for side quality

quincunx sub-sampling method is chosen, but 2-pixels est. and 4-pixels est. will be both used.

For all the cases, we use the video coding standard H.264 [10][11] as the basic coder. H.264 is the latest video coding standard and also known as MPEG-4 part-10 (AVC). It changes a lot compared to previous version H.263, such as 4 by 4 DCT and much more coding modes of a macro block, etc., which make it better than H.263 for 3-5dB. Also arbitrary frame size is supported and this gives much more flexibility to implement our approach which may add different number of zeros.

Fixed frame rate (30frames/second) and constant quantizer step size are used for all sequences. Note that the frame rate doesn't affect the estimation results much because all the estimation is based on within current frame, not using other adjacent frames. No B frame is used. Entropy coding is CAVLC. Two sequences are used of which one is 'paris' of CIF and the other is 'foreman' of QCIF.

Fig. 2 shows the performance of our 1D-MDC approach and 2D-MDC for various numbers of padded zeros using 'paris' CIF sequence. We consider combined bit rate vs. Central PSNR and Side PSNR. Combined bit rate is the sum of bit rates of two descriptions. Central/Side PSNR is the average PSNR of reconstructed video for luminance. One common feature of figures (a)-(d) is that performance will be better with increasing zeros, except that for very high bit rate, the central decoding quality of more zero padding is a little worse. The reason is that padding more zeros (which means higher correlations) decreases bit rate by making more zeros after quantizer and less coefficients to be coded, while quantizer with very small QP generates many more small coefficients and this results in low efficiency for padding more zeros. In above four figures, for every cluster of curves of one method, number of zeros is 32, 96, 192, 288 respectively from bottom to above. Fig. 2 (a) compares central quality of the Interleaved and Quincunx sub sampling methods for 1D-MDC. We can see that for interleaved sub

sampling, there is not so much difference between various zero padding, and the difference is a little clearer for Quincunx sub sampling due to much more zeros padded. Fig. 2 (b) shows the side decoding quality of the estimation methods using quincunx 2-pixels est. and quincunx weighted 4-pixels est. for 1D-MDC. The weight increases with more zeros, and at last it reaches 1. Quincunx weighted estimation method is better than quincunx 2-pixels estimation method for small number of padded zeros, and they will be equal for large number of zeros. From Fig. 2 (c) it's shown that interleaved method is always better than quincunx method except for small number of zeros with high bit rate. In Fig. 2 (d) Side decoding quality of estimation method using 2-pixels est. and 4-pixels est. for 2D-MDC are compared. For certain number of zeros padded, 4-pixels mean is always better than just column based mean. This advantage decreases with more zeros padded. But for higher zero padding, central quality of 2D-MDC is not good because so many zeros are added in. Experiments show that around 160 and 192 zeros is enough for it. This means that we cannot add more zeros for higher side PSNR if we want a good central decoding quality in 2D-MDC.

For our 1D-MDC, adding more zeros makes both the central and side quality better. It brings much better performance over 2D-MDC which is shown in next figures. Next we give the comparison between the best modes of 1D-MDC and 2D-MDC. Fig. 3 illustrates the difference of them.

It's clear from Fig. 3 that interleaved 1D-MDC is much better than 2D-MDC for both central and side quality. There are mainly two reasons. One is that 2D-MDC adds so many zeros which decreases efficiency especially for low bit rate. The other is that, because 2D-MDC pads zeros in each direction, it's more suitable to use quincunx sub sampling method. But for this application, quincunx sub sampling decreases correlations between pixels and adds more high frequency in image, which results in low efficiency as can be seen in both comparisons for 1D-MDC and 2D-MDC (Fig. 2 (c) and Fig. 3).

It should be noted that there is a limit for number of padded zeros. Adding zeros will increase the correlation between pixels while also increase the picture size. When the increased correlation cannot counteract the effect of bigger size, the performance will drop sharply with increasing zeros. Moreover, increasing zeros will bring mismatch in itself because of removing decimal fraction part as mentioned above. Because the total energy of picture doesn't change, the effect of fraction must be more and more important with increasing zeros padded. If we add enough zeros on the picture, values of pixels will be all below 0.5, which means there will be nothing in the picture after the pre-processing stage.

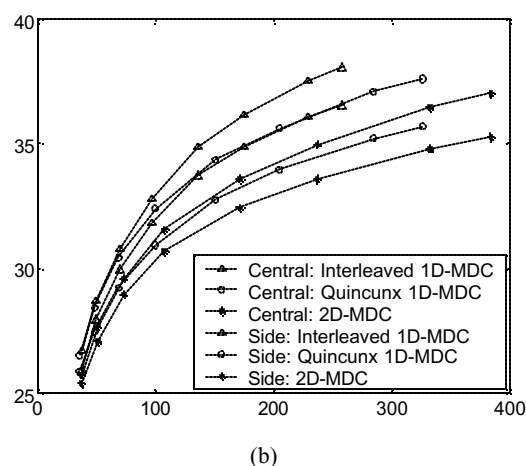
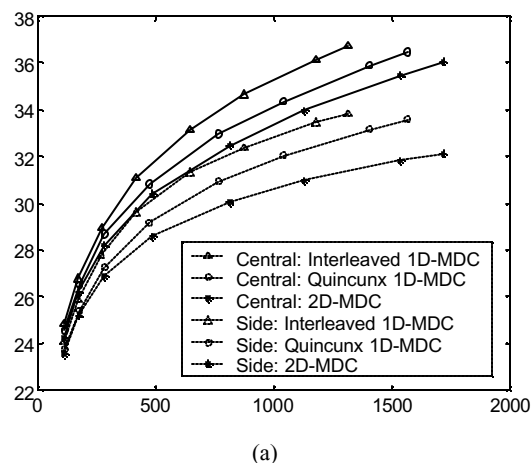
In addition to outperform 2D-MDC in both central and side quality, 1D-MDC also reduces computation complexity. 2D DCT is separable and can be thought of as two 1-D DCT. Moreover, the size of IDCT in 2D-MDC is larger than 1D-MDC because 2D-MDC pads much more zeros. So if we calculate the complexity of the pre-processing stage, 2D-MDC is at least two times of that of 1D-MDC.

#### 4. CONCLUSIONS

In this paper we introduce one MDC approach based on one dimensional zero padding, which improves the performance and reduce the complexity of [9]. It can be completely implemented through pre- and post-processing. We add redundancy by padding various numbers of zeros in one dimension DCT domain. Multiple descriptions are generated by sub sampling zero padded frames. It's shown through simulations that 1D-MDC using interleaved sub sampling method performs much better than 2D zero padding method, and for 1D-MDC, interleaved sub-sampling method is better than quincunx sub-sampling

#### REFERENCES

- [1] V. Vaishampayan, "Design of multiple description scalar quantizers," IEEE Trans. On information Theory, vol. 39, no.3, pp.821-834, May 1993.
- [2] V. Vaishampayan and S. John, "Interframe balanced-multiple-description video compression," Packet Video '99, New York, NY, USA, Apr. 1999.
- [3] M. Orchard, Y. Wang, V. Vaishampayan, and A. Reibman, "Redundancy rate distortion analysis of multiple description coding using pairwise correlating transforms," International Conference on Image Processing, Santa Barbara, CA, USA, Oct. 1997.
- [4] Y. Wang, M. Orchard, and A. Reibman, "Optimal pairwise correlating transforms for multiple description coding," International Conference on Image Processing, Chicago, Illinois, USA, Oct. 1998.
- [5] A. Reibman, H. Jafakhani, Y. Wang, and M. Orchard, "Multiple description coding for video using motion compensated prediction," International Conference on Image Processing, Kobe, Japan, Oct. 1999.
- [6] J. G. Apostolopoulos, "Error-resilient video compression through the use of multiple states," Proc. IEEE Int. Conf. Image Proc, Vancouver, CA, USA, Sept. 2000.
- [7] C. -S. Kim and S. -U. Lee, "Multiple description motion coding algorithm for robust video transmission," IEEE Int. Symp. on Circuits and Syst., Geneva, Switzerland, May 2000.



**Fig. 3 Comparison between 1-D MDC and 2-D MDC: Bitrate (kbps) vs. PSNR (dB).**

(a) 'paris' CIF (b) 'foreman' QCIF

- [8] S. Shirani, M. Gallant and F. Kossentini, "Multiple Description Video Coding Using Pre- and Post-Processing," International TCC, Las Vegas, Nevada, USA, April 2001.
- [9] M. Gallant, S. Shirani, F. Kossentini, "Standard-compliant Multiple Description Video Coding," 2001 IEEE International Conference on Image Processing Thessaloniki, Greece, October 2001.
- [10] H.264 standard, JVT-G050, 7th meeting, Pattaya, Thailand, 7-14 March, 2003
- [11] Thomas Wiegand, Gary J. Sullivan, Gisle Bjontegaard, and Ajay Luthra, "Overview of the H.264/AVC Video Coding Standard," IEEE Transactions on Circuits and Systems for Video Technology, July 2003.
- [12] G. Olmo, T. Tillo, "Directional Multiple Description Scheme for still images," Submitted to 10th IEEE International Conference on Electronics, Circuits and Systems, April 2003.